

#6

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/557,796

DATE: 10/20/2000
 TIME: 11:48:15

Input Set : A:\557796.txt
 Output Set: N:\CRF3\10202000\I557796.raw

4 <110> APPLICANT: Hoch, James
 5 Dartois, Veronique
 7 <120> TITLE OF INVENTION: METABOLIC SELECTION METHODS
 10 <130> FILE REFERENCE: 234/191
 12 <140> CURRENT APPLICATION NUMBER: 09/557,796
 13 <141> CURRENT FILING DATE: 2000-04-25
 15 <150> PRIOR APPLICATION NUMBER: 09/172,952
 16 <151> PRIOR FILING DATE: 1998-10-14
 18 <160> NUMBER OF SEQ ID NOS: 33
 20 <170> SOFTWARE: FastSEQ for Windows Version 3.0
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 816
 24 <212> TYPE: DNA
 25 <213> ORGANISM: yia j
 27 <400> SEQUENCE: 1
 29 atgggcacaa aagaaagcga gaacacgcaa gataaagaga ggcctgcccg aagtcagagc 60
 30 ctttttcgtg ggttgatgct aattgagatc ctgagtaatt atccaaatgg ctgtcccgtg 120
 31 gcgcattctg cggaaactggc gggactgaac aaaagtaccg ttcactcgctt attacagggg 180
 32 ctgcagtcct gcgggtacgt gacgcctgcc ccggcggcgg ggagctatgc gctgacgaca 240
 33 aaatttatcc gcgttgccca aaaggcgttg tcgtcgctga atattatcca cgtcgcggcg 300
 34 ccgcattctg aggcgcttaa cctggccacc ggcgagacgg tgaacttctc cagccgtgaa 360
 35 gatgaccacg cgatccctgat ttataagctg gagccgacca ccggtatgct gcgtacgcgc 420
 36 gccttatatt gccagcacat gcgctgtact gctcggcaat gggcaaagat ttatatggcg 480
 37 tttggccatc ctgactacgt tgagagctac tggaattcac accaggagat tatccagccg 540
 38 ctgaccgcga ataccattac cggcttgctc gcgatgcag atgaactggc gcagatccgc 600
 39 gaggcgaata tggcgatgga cagggaagag aacgagctgg gcgtgtcgtg cctggctgtc 660
 40 ccgcgttttg atatccatgg gcgcgtgcct tatgccattt ctatctctct atcaacatcg 720
 41 cgcctcaagc aggtgggaga gaaaaattta ctcaagccgc tacgcgatac ggcagagggc 780
 42 atttctcgcg aactgggctt ttccgtgcgg gaaggt 816
 45 <210> SEQ ID NO: 2
 46 <211> LENGTH: 996
 47 <212> TYPE: DNA
 48 <213> ORGANISM: yia k
 50 <400> SEQUENCE: 2
 52 atgaaagtca cgtttgagca gttaaaagag gcattcaatc gggtagctgt ggacgcgtgc 60
 53 gtcgcccggg aaaccgcca tgctgcgca gaaatgtttg ccgcaccac cgaatccggc 120
 54 gtctattctc acggcgtaga ccgctttcct cgcttcattc agcagttgga taacggcgac 180
 55 attatccctg aggtcaacc gcagcgggtg accacgctcg gcgccatcga acagtgggat 240
 56 gctcagcgtt ccatcggaac cctgacggcg aaaaagatga tggatcgggc cattgagctg 300
 57 gcctccgatc acggtatcgg cctggctgcc ttacgtaatg ctaaccactg gatgcgcggc 360
 58 ggcagctacg gctggcaggc ggcggaaaaa ggctacatcg gtatctgctg gaccaactcc 420
 59 atcgcggtta tggcgccatg ggcgcgtaaa gagtgcctga tcggtacca cccgctgata 480
 60 gtcgccaatc cgtcgacgcc gatcaccatg gtggatatgt cgatgtcgat gttctcctac 540
 61 ggcattgctg aggttaaccg ccttgccggc cgcgaactgc ccgtggacgg cggattcgac 600
 62 gatgacggtc gtttgaccaa agagccgggg acgatcgaga aaaatcgccg cattttaccc 660
 63 atgggctact ggaaagggtc cggcctgtcg atcgtgctgg atatgattgc caccctcctc 720
 64 tccaacggat cgtcggttgc cgaagtgaac caggaaaaa cegatgaata tggcgtttcg 780

ENTERED

RAW SEQUENCE LISTING

DATE: 10/20/2000

PATENT APPLICATION: US/09/557,796

TIME: 11:48:15

Input Set : A:\557796.txt

Output Set: N:\CRF3\10202000\I557796.raw

```

65  cagatcttca tcgctattga agtggaataag ctgatcgacg gcgcaaccgc cgacgccaag      840
66  ctgcaacgga ttatggattt catcaccacc gccgagcgcg ccgatgaaaa tgtggcggtc      900
67  cgtcttcctg gccatgaatt tacccgtctg ctggatgaaa accgccgcaa cgccattacc      960
68  gtcgatgaca gcgatggggc caaaattcag gcgctg                                     996
71  <210> SEQ ID NO: 3
72  <211> LENGTH: 462
73  <212> TYPE: DNA
74  <213> ORGANISM: yia.1
76  <400> SEQUENCE: 3
78  atgatttttg gtcattatgc tcaacctaat ccgtgtcgtc tgcccgcggc cattgagcgg      60
79  gcgcttgatt tcctgcgcac gacggatttc cagcgctggg caccgcgctg cgtggaaatc      120
80  gacggccaac acatcttcgc gcaggttatc gacttaacca ctgcgatgc cgtgaaaaat      180
81  cgtccggagg tcaccgctgc ctatctggat atccagtttc tggcatcggg cgaagaaaaa      240
82  atcggtatcg ccattgatac cggcaataat caaatcagcg aatctttatt agaacagcgc      300
83  gatattatct tttatcacga cagcgaacat gaatcgttct ttgaaatgac gccaggcaac      360
84  tatgcgatat tttcccgca agatgttcat cgtcctggat gtaataaaac tgtagccacg      420
85  ccgatccgca aaatagtcgt taaagtcgct atttcagttt ta                                     462
88  <210> SEQ ID NO: 4
89  <211> LENGTH: 945
90  <212> TYPE: DNA
91  <213> ORGANISM: orf1
93  <400> SEQUENCE: 4
95  atgaattcga ataataccgg ttacattatc ggtgcgtacc cctgtgcccc ctgtgcaccc      60
96  tcatttcacc aaaagagtga agaggaagag atggaattct ggcggcagct ctccgacacc      120
97  ccggatattc gcgggctgga gcaaccctgc ctaccctgcc ttgaacatct tcattccgctc      180
98  ggcgacgagt ggttattgag ccataccccc ggacactggc agattgtcgt tacccgcatc      240
99  atggaaacca tgcgccgccc cggtgaaaac ggcggctttg ggctggcgct cagcgacgaa      300
100 acgcagcgca aagcctgcgt ggagtactat cgccacctgc agcagaagat cgctaaaatc      360
101 aatggcaata ccgccggaac ggtcattgcc cttgagcttc acgccgcccc gctggcgggc      420
102 aatgccaacg tggctcaggc tacgcagcgc tttgcccggt cattaaaaga aattaccgcg      480
103 tgggactggt cctgcgagct ggtgctggag cactgcgacg cgatgaccgg cagcgcgccg      540
104 cgcaaggat ttttgccgtt agaaaaacgt ctggaagcca ttgcccatta tgacgttggc      600
105 atttgtatta actggcgcg cttcgccatt gaaggcgcca ataccgtgct accgctcacc      660
106 catacgagc aggtaaaacg ggcaggaaag ctgcggcgcg tgatgttttc tggcacgacg      720
107 cagaccggcg agtacggcga atggcaggat ttacacgcgc cgttcgcgcc tttctgcccc      780
108 cagagcctga tgaccaccga acacgctcgt gaattatttg cctgcgcagg aaccgcccc      840
109 ctgcaatttt caggcattaa attactggaa attaatgcca gcgcaaacgt tgatcatcgc      900
110 atcgcgatat tacgcgacg catctccgcg ctaaaacaag caciaa                                     945
113 <210> SEQ ID NO: 5
114 <211> LENGTH: 1317
115 <212> TYPE: DNA
116 <213> ORGANISM: yia x2
118 <400> SEQUENCE: 5
120 atgaatataa cctctaactc tacaacaaaa gatataccgc gccagcgctg gttaagaatc      60
121 attccgccta tactgatcac ttgtattatt tcttatatgg accgggtcaa tattgccttt      120
122 gcgatgcccg gaggtatgga tgccgactta ggtatttcgg ccaccatggc ggggctggcg      180
123 ggcggtattt tctttatcgg ttatctatct ttacaggttc ccggcgggaa aattgcccgt      240
124 caggtagcgc gtaagaaatt tatcgctggg tcgctggctg cctggcggtt catctccgtg      300
125 ctgacggggg taattaccaa tcagtaccag ctgctggccc tgcgcttctt actggcgctg      360

```

RAW SEQUENCE LISTING

DATE: 10/20/2000

PATENT APPLICATION: US/09/557,796

TIME: 11:48:15

Input Set : A:\557796.txt

Output Set : N:\CRF3\10202000\I557796.raw

```

126 gcggaaggcg gtagtgcgcc ggtcggtctc acgatgatca gtaactgggt ccccgacgct 420
127 gaacgcgggc gcgccaacgc gattgtcatt atgtttgtgc cgattgccgg gattatcacc 480
128 gcccactctc caggctggat taccacggtt ctcgactggc gctggctgtt tattatcgaa 540
129 ggtttgctct cgtcggttgt tctggttctg tgggcataca ccatctatga ccgtccgcag 600
130 gaagcgcgct ggatttccga agcagagaag cgtatctcgg tcgagacgct ggccgcggag 660
131 caaaaagcca ttgccggcac cgaggtgaaa aacgcctctc tgagcgccgt tctctccgac 720
132 aaaaccatgt ggcagcttat cgcctgaac ttcttctacc agaccggcat ttacggctac 780
133 accctgtggc taccaccat tctgaaaaga ttgaccata gcagcatggg gcaggtcggc 840
134 atgettgcga ttctgcgta cgtcggcgcc attgctggga tgttctctgt ttcctccctt 900
135 tcagaccgaa ccggttaacg caagctgttc gtctgcctgc cgtcgattgg ctctcgtctg 960
136 tgcattgtcc tgtcgggtgg gctgaaaaac caaatttggc tctctatgc cgcgctgggtc 1020
137 ggctcgggat tcttctcgca atcggcggtt ggcgtgttct ggaccatccc ggcacgtctg 1080
138 ttcagcgcgg aaatggcggg cgcgcgcgcg ggggttatca acgcgcttgg caacctcggc 1140
139 ggattttgtg gcccttatgc ggtcgggggt ctgacacgt tgtacagcaa agacgctggc 1200
140 gtctattgcc tggcgatctc cctggcgctg gccgcgctga tggcgctgct gctgcggcg 1260
141 aaatgcgatg ccggtctcgc gccggtaaaag acgataaacc cacataaacc cactgcg 1317
144 <210> SEQ ID NO: 6
145 <211> LENGTH: 1503
146 <212> TYPE: DNA
147 <213> ORGANISM: lyxk
149 <400> SEQUENCE: 6
151 atgagcaaga aacaggcctt ctggctgggt attgattgcg gcggcaccta tctgaaagcc 60
152 ggtttatatg acgcgaagg tcatgaacat ggcattgtgc ggcaagcgct acggacgatg 120
153 tcgcccctgc cgggttacgc cgaacgcgac atgcgccagc tctggcaaca ctgcgcggcg 180
154 accattgccc ggctattaca gcaggcaggt gtatccggcg aacagattaa aggcgtgggc 240
155 atctcgcctc agggtaagg gctcttctc ctcgataagc aggatcggcc gctgggtaac 300
156 gccatcctct cctccgatcg tcggcgctcg aaaatcgctc agcgtggca gcgggaccgt 360
157 attcccgaa cgccttatcc cgttaccgcg cagacgctgt ggaccggaca tccggttct 420
158 ttgtcgcgct gggtaaaaga gaatgaaccc cagcgtctac gcgaaattgg ctgcgtgatg 480
159 atggggcatg actatctgcy ctggtgctta accggcgcg aaggctgcga ggagagcaac 540
160 atctccagat ccaacctcta caacatggcc atgggccagt acgacccgcg cctgaccgag 600
161 tggctgggca tcgggtgaaat ctagcgcg cgtcccccg ttgtagggtc agccgaaatt 660
162 tgcggggaga tcaccgctca ggcagccgct ttaaccggtc tggcgcgggg tactccgctc 720
163 gttggcggcc tgtttgacgt ggtctccacc gccctttcgc ccgggattga ggatgagtcg 780
164 accctcaatg cgggtgatgg gacctgggcc gtcactagcg gtatcgctca cggcctgcgc 840
165 gaccatgagg cccaccctta cgtctatggc cgtacgtca atgacggcca gtatatcgtt 900
166 caggaagcca gcccgacctc atccggcaac ctcgaatggt ttaccgcca gtggggcgat 960
167 ctctcgtttg atgagatcaa tcaggcgcgt gccagcctgc cgaaagccgg gagcgagctg 1020
168 ttttttctgc cgtttctgta tggcagcaac gccgggctgg agatgacctg cggcttttac 1080
169 ggcatgcagg cgtgcatac ccgcgcgcac ctgctgcagg cggtttatga aggcgtggta 1140
170 tttagccata tgaccacact cagccgtatg cgcgaacgct ttacaaacgt tcaggccctg 1200
171 cgcgtcacgg gcggcccggc gcaactcgac gtcgtgatgc agatgctggc ggacgtaagc 1260
172 ggccttacga ttgaactccc gaaggtggaa gagaccggct gttttggcgc ggccctcgcc 1320
173 gctcgtgtcg gtaccggcgt ataccgcagc tttagcgaag cccggcgcg cggcgagcac 1380
174 ccggtgcgca cgtgctgccc cgatatgacc gccacgcgc gctatcagcg caaataccgc 1440
175 cactacctgc atttgattga agcactacag ggctatcacg cccgtattaa ggagcacgca 1500
176 tta
179 <210> SEQ ID NO: 7
180 <211> LENGTH: 660

```

RAW SEQUENCE LISTING DATE: 10/20/2000
 PATENT APPLICATION: US/09/557,796 TIME: 11:48:15

Input Set : A:\557796.txt
 Output Set: N:\CRF3\10202000\I557796.raw

```

181 <212> TYPE: DNA
182 <213> ORGANISM: yia q
184 <400> SEQUENCE: 7
186 atgagccgac cattaactgca gctggcgctc gaccatacca gccttcaggc tgcgcagcgc 60
187 gatgtcgcgc tgctacagga tcacgttgat attgtggagg cgggaacct cctctgctta 120
188 accgaagggc ttagcgcggt taaagccctg cgcgcccagt gtccggggaa gatcatcgtc 180
189 gccgactgga aagtcgccga cgcggtgtaa accctggcgc agcaggcctt tggcgctggc 240
190 gccaaactgga tgacctcat ttgcgcgcga ccgctcgcca cggctcgaga aggccacgcc 300
191 gtggcccagg cctgcggcgg tgaattcag atggagctgt tcggcaactg gacgctggat 360
192 gacgcccgcg cctggtaccg taccggcgtc catcaggcga ttaccatcg cggacgcgat 420
193 gcccaggcca gcgggcagca gtggggggag gcgcatctgg cgcgcatgaa agcgtgttcc 480
194 gatattggcc ttgagctatc gattaccggc ggcattaccc cagccgatct accgctgttc 540
195 aaagatatca acgtcaaaag ctttattgcc gggcgcgcg tggaaggcgc cggccatccg 600
196 gcgcgggttg ccgcggaatt ccacgcgcaa atcgacgcta tctggggaga acagcatgcg 660
199 <210> SEQ ID NO: 8
200 <211> LENGTH: 858
201 <212> TYPE: DNA
202 <213> ORGANISM: yia r
204 <400> SEQUENCE: 8
206 atgcgtaacc acccggttagg tatttatgaa aaagcgctgg cgaaggatct cagctggcct 60
207 gageggctgg tactggccaa aagctgcggt tttgattttg tcgaaatgtc ggtggacgag 120
208 accgatgaac gccttttcgc cctggagtgg acccgggccc agcgcgcacg gctggtgagc 180
209 gcgatgctgg aaaccgcggt cgcattcccc tcgatgtgct tgccgcacca tcgcgctttc 240
210 ccttttgga gccgcgatga agcggtagcg gatcgggcgc gagagattat gaccaaagcc 300
211 atccgcctgg cgcgcgatct ggggatccgc accatccagc tggcgggtta cgacgtctat 360
212 tacgaagagc atgatgaagg caccggcgag cgttttgccg aagggctggc ctggggcggt 420
213 gaacaggccg ccgcccgcga ggtaatgctg gcggtggaga tcatggacac cgcctttatg 480
214 aactccatca gcaaatggaa aaagtgggac gagatgctt cgtaccctg gttaccgtc 540
215 taccggagc tcggcaacct cagcgcctgg ggaacgcag tcaccgccga gctgaagctg 600
216 ggcacgcgac gtatcgccgc catccacctg aaagatacgc tgcccgtgac cgacgatagc 660
217 cctggccagt tccgcgacgt gccgttcggc gaaggatcgc tcgattttgt cggcattttt 720
218 aagacgctgc gcgagctgaa ctaccgcggt tcatttttga ttgagatgtg gacggagaaa 780
219 gccagcgcgc cgggtgctgga gattatccag gcccggcgct ggatcgaatc acggatgcag 840
220 gaagggggat tcacatgt 858
223 <210> SEQ ID NO: 9
224 <211> LENGTH: 714
225 <212> TYPE: DNA
226 <213> ORGANISM: yia s
228 <400> SEQUENCE: 9
230 atgttagaac aactgaaagc cgaggtagct ggcgcaaacc tggccctccc cgcacacggc 60
231 ctggtcacct ttacctgggg caacgtcagc gcggtcgatg aaacgcgcaa gctgatggtc 120
232 attaaacctt ccggcgtcga atatgaggtg atgaccgcgc acgatatggt ggtcgtagag 180
233 atggccagcg gtaaaagctg tgaaggcggg aaaaaacctt cttcagatac gccaacgcac 240
234 ctggcgcttt atcgccgcta tccgcagatc ggcgggatcg tgcataccca ctcccgccac 300
235 gcgacgatct ggtcgcaggc cgggctcgat ctccccgcct ggggcaccac ccacgccgac 360
236 tactttctatg gcgcgatccc ctgtaccgga cggatgacgc ttgaggagat taacggcgag 420
237 tatgagtatc agaccggcga ggtgattatc aaaacctttg aacagcgcg cctggatccg 480
238 gcgcaaatcc cggcggtatt ggtccattca cagggccctt ttgctgggg taaagacgcc 540
239 gccgacgcgc tacataacgc cgtggtgctg gaggagtgcg cctacatggg cctcttctcg 600

```

RAW SEQUENCE LISTING DATE: 10/20/2000
 PATENT APPLICATION: US/09/557,796 TIME: 11:48:15

Input Set : A:\557796.txt
 Output Set: N:\CRF3\10202000\I557796.raw

```

240 cgccagtggc cacagctgcc ggatatgcag tctgaactgc tcgataaaca ctatctgcgt      660
241 aaacacggcg cgaacgctat tacgggcaaa actagtcccg cggaactccc cgga      714
244 <210> SEQ ID NO: 10
245 <211> LENGTH: 272
246 <212> TYPE: PRT
247 <213> ORGANISM: YiaJ-Ko
249 <400> SEQUENCE: 10
251 Met Gly Thr Lys Glu Ser Glu Asn Thr Gln Asp Lys Glu Arg Pro Ala
252 1 5 10 15
254 Gly Ser Gln Ser Leu Phe Arg Gly Leu Met Leu Ile Glu Ile Leu Ser
255 20 25 30
257 Asn Tyr Pro Asn Gly Cys Pro Val Ala His Leu Ser Glu Leu Ala Gly
258 35 40 45
260 Leu Asn Lys Ser Thr Val His Arg Leu Leu Gln Gly Leu Gln Ser Cys
261 50 55 60
263 Gly Tyr Val Thr Pro Ala Pro Ala Ala Gly Ser Tyr Ala Leu Thr Thr
264 65 70 75 80
266 Lys Phe Ile Arg Val Gly Gln Lys Ala Leu Ser Ser Leu Asn Ile Ile
267 85 90 95
269 His Val Ala Ala Pro His Leu Glu Ala Leu Asn Leu Ala Thr Gly Glu
270 100 105 110
272 Thr Val Asn Phe Ser Ser Arg Glu Asp Asp His Ala Ile Leu Ile Tyr
273 115 120 125
275 Lys Leu Glu Pro Thr Thr Gly Met Leu Arg Thr Arg Ala Tyr Ile Gly
276 130 135 140
278 Gln His Met Arg Cys Thr Ala Arg Gln Trp Ala Lys Ile Tyr Met Ala
279 145 150 155 160
281 Phe Gly His Pro Asp Tyr Val Glu Ser Tyr Trp Asn Ser His Gln Glu
282 165 170 175
284 Ile Ile Gln Pro Leu Thr Arg Asn Thr Ile Thr Gly Leu Pro Ala Met
285 180 185 190
287 His Asp Glu Leu Ala Gln Ile Arg Glu Arg Asn Met Ala Met Asp Arg
288 195 200 205
289 Glu Glu Asn Glu Leu Gly Val Ser Cys Leu Ala Val Pro Val Phe Asp
290 210 215 220
292 Ile His Gly Arg Val Pro Tyr Ala Ile Ser Ile Ser Leu Ser Thr Ser
293 225 230 235 240
295 Arg Leu Lys Gln Val Gly Glu Lys Asn Leu Leu Lys Pro Leu Arg Asp
296 245 250 255
298 Thr Ala Glu Ala Ile Ser Arg Glu Leu Gly Phe Ser Val Arg Glu Gly
299 260 265 270
303 <210> SEQ ID NO: 11
304 <211> LENGTH: 332
305 <212> TYPE: PRT
306 <213> ORGANISM: YiaK-Ko
308 <400> SEQUENCE: 11
310 Met Lys Val Thr Phe Glu Gln Leu Lys Glu Ala Phe Asn Arg Val Leu
311 1 5 10 15
313 Leu Asp Ala Cys Val Ala Arg Glu Thr Ala Asp Ala Cys Ala Glu Met

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/557,796

DATE: 10/20/2000

TIME: 11:48:16

Input Set : A:\557796.txt

Output Set: N:\CRF3\10202000\I557796.raw